8. Simulation Philosophy and Methods

Steve Harrison

In general, to 'simulate' means to mimic or capture the essence of something, without attaining reality. For management-oriented applications, the something is an identified *system* under the control of management – for example of a *bioeconomic* system – and the essence is captured by way of a symbolic or algebraic *model*. Simulation consists of the steps of developing the model to represent a *real system*, and then performing *experiments* using this model to predict how the real system would behave under a range of *management policies*. The objective of simulation may be to increase understanding of the behaviour of the system, or to compare various policies for management of the system. While many quantitative techniques take a well-recognized form, simulation differs in its great flexibility, variety of applications and variations in form. These features, while highly valuable for modelling complex systems, make this a difficult methodology to explain and to comprehend. In fact, simulation has been described as 'more art than science'. Proficiency with this technique cannot be gained easily in the classroom, but rather requires considerable practical experience. This module presents the basic concepts of simulation and the steps fundamental to simulation studies. The module first defines the nature of simulation and the philosophy behind this technique. Modelling concepts and elements of models are then discussed. The typical steps when using simulation are next outlined. A simple example of developing and applying a model is presented to aid discussion. Some comments are made about validation of simulation models and about design of simulation experiments.

1. WHAT IS SIMULATION?

Simulation consists of a number of disparate concepts and techniques, with different terminologies adopted by different disciplines. Hence it is useful to establish the terminology which will be adopted in this discussion of the application of simulation to bioeconomic systems.

Proponents of simulation usually subscribe to what is called the *systems approach*. According to Shannon (1975), a system is 'a group of objects united by some form of interaction or interdependence to perform a specified function'. In other words, any system consists of a number of *interrelated* and *interacting* parts; further, these parts should not be studied in isolation but rather in the context of the overall system and its complex interdependencies. The whole is more than just the sum of the parts. Any change to one part of the system may cause unexpected changes elsewhere. Scientists, engineers, managers and so on have increasingly realized that it is necessary to take this *holistic view* of the systems they design and operate.

As the body of scientific knowledge has increased, there has been a tendency for greater specialization of research, with a loss in overall perspective and loss of communication between researchers in different disciplines. This 'spread of specialized deafness' has led to the study of more narrowly defined systems. This is not to suggest that *reductionist* research into narrowly defined systems does not play an important role in advancement of knowledge. However, from a management point of view, the system of interest is usually the level of aggregation at which planning and control decisions are made, which is often an overall business firm or a particular project being undertaken by a firm.

Various terminologies have been adopted in presentations of simulation methodology. In particular, the terms *systems analysis*, *systems research* and simulation have been used interchangeably. The term 'systems research' is typically applied to describe all the steps in the study of an organized system. 'Systems analysis' was originally used in this broad context, but has recently become applied more often to just one step in systems research, viz. that of identifying the boundaries, elements and interrelationships of a system prior to modelling it. The term 'simulation' is sometimes applied to the overall procedure of modelling and experimentation, and sometimes to the experimentation stage only. In this module the term 'simulation' will be used in the broader context, as a synonym for systems research.

Under the systems approach, a model is developed which represents as closely as possible the essential structure and performance of the real system, in terms of the specific behavioural features the researcher wishes to examine. This model once constructed and tested for reliability is then used to simulate or mimic how the actual system would behave under particular circumstances, by conducting experiments on a computer using the model. In these experiments, measures of system performance are generated when levels decision variables are set at a number of levels.

Simulation tends to be used where a solvable model is not available (i.e. would not represent the complex structure of the real system adequately). The simulation experiments may be likened to observing how the real system would perform if particular management policies were to be adopted, except that the real system is not interfered with, real resources are not used (apart from computing resources), and time is greatly compressed. Computer simulation experiments can thus provide a great deal of information about how an actual system would behave, under a host of different policies and environments, and this may provide a greatly improved understanding of the system and how to manage it.[1]

Practitioners of the systems approach need to have a good understanding of the overall system, and the ability and willingness to consult with experts on various components of the system. In fact, systems research is often conducted by groups or *multidisciplinary teams* rather than individuals, since these can take on board a broader range of expertise on various aspects of the system.[2]

2. THE NATURE OF MODELS

The term 'model' has wide everyday use. Physical or *iconic* models are typically scaled down versions of real systems. Some physical models are not true-to-scale replicas, but still convey information about a system. In a broad sense, any map is a model of a territory, and any timetable is a model of the operation of a transport system. Hence, any form of model building could be thought of as a form of simulation. However, the term has become associated with a particular approach to decision support. From a forestry research perspective, the models we are interested in are usually algebraic models of bioeconomic systems.

Nowadays, people are familiar with building algebraic models, through the widespread use of spreadsheets. Any spreadsheet – including one to derive the net present value (NPV) of a project – is in effect an algebraic model. The spreadsheet allows experimentation with the model, such as asking 'what if' type questions through varying a model input and checking the performance estimates.

---

[1] A particular type of simulation is that of various forms of *games*. There is a long history of war games, both using model ships or soldiers on a table, or using actual troops but dummy ammunition. In recent years, enormous resources have been put into development of computer games, and some of the computer graphics in these have become incredibly good simulations of real world scenes and activities. Another gaming application is that of *management games* for training in business, and macroeconomic games for training in economics. For these games to be considered genuine by participants, a credible model must be present which mimics the behaviour of the underlying business or economic system for which the training is being conducted.

[2] The alternative is to have studies conducted by *transdisciplinary* individuals, i.e. researchers with multiple skills across a range of disciplines. The obvious difficulty here is for individuals to have sufficient depth across a range of disciplines to adopt a systems approach.

The need for a model

Sometimes it is possible to gain experience about the operation of a system by imposing various managements upon the system itself. Agricultural researchers have traditionally used research stations and on-farm trials to evaluate pasture species, cropping systems, irrigation practices and so on. Many business managers would admit to learning how to operate their business in the ‘school of hard knocks’. However, learning by one’s mistakes can be a slow, costly and unpredictable way of gaining experience. Conducting experiments on a model rather than the real system may be preferable on a number of grounds:

(i) the real system might not yet exist. For example, a company may wish to expand or diversify its activities, perhaps investing a large amount of capital. It would be useful to have an idea of potential outcomes of this venture, and their likelihoods, before committing capital to it.

(ii) performing experiments on the real system may be unacceptably time consuming or expensive. This would be the case, for example, if management wished to know the cash flows resulting over the next 10 years if forestry plantations were to be located in either of two alternative sites or either of two different species or mixtures were adopted. Once a model is constructed, stand growth and economic performance can be generated and compared on the computer for the two sites or species, in a matter of seconds of real time.

(iii) performing experiments may be inimical to the real system. For example, suppose the directors of a forestry company wished to know whether the company could survive under various levels of borrowed finance. It would certainly be a risky proposition for the company to take out a very large loan and see if the debt could be serviced. However, no harm would be done to the company if a model were used to experiment with different gearing ratios, and to determine what debt level leads to an acceptably low probability of ‘simular’ bankrupcy or takeover.

(iv) The exercise of developing a model can be in itself a valuable learning experience about a system. A model provides a framework for systematically organizing existing information, and for assembling the (often subjective) knowledge of experts. The explicit nature of models makes ideas and assumptions more transparent, and places these under the scrutiny of more people. This in part explains why models tend to be built through a series of prototypes, evolving over time as questionable assumptions and gaps in knowledge are exposed, more questions are asked and more information is assembled.

Types of models

Typically, a bioeconomic system is modeled as a set of equations or relationships between variables, although often the algebra of the model is implied rather than made explicit. Various classifications of symbolic models have been advanced. For example, models may be grouped as

(i) biological, economic or bioeconomic
(ii) discrete or continuous
(iii) single-period or multiperiod (or static or dynamic)
(iv) deterministic or stochastic.

A great deal of modelling has been carried out in the natural sciences, including forestry. For example, models have been developed of physiological processes within trees, and of the growth of individual trees and stands of trees. A *biological* model of a forestry stand could take into account soil moisture, nutrient uptake, biomass growth, partitioning of nutrients between plant parts, and so on. The time step would normally be a period less than a year, perhaps as short as a day. The performance measure might be total biomass or bole biomass produced. A *financial* or *economic* model of a forestry stand would typically take account of expenditure and of product sales, on a yearly basis, and seek to estimate the net present value of the forestry investment.[3] Components of a biological and an economic model may be combined to

[3] Some differences exist between financial and economic models, as discussed in Modules 15 and 16.

produce a *bioeconomic* model, which predicts profitability from the forestry investment, but taking into account the production processes rather than simply taking final yield estimates.

Models in which time changes in a *continuous* fashion are used extensively in mathematics and engineering, frequently taking the form of sets of simultaneous differential equations. Most management-oriented models are of *discrete* form, with variables taking time steps of say one week or one year.

Static models make no allowance for changes in values of variables over time. They are thus not suited to examining long-term production processes such as forestry systems. Multiperiod or dynamic models include specific representation of changes in levels of variables over time, such as growth, partition and decay of biological components and accumulation of funds or assets.

In a *deterministic* model, all variables take single point or best-estimate values, and all relationships between variables are assumed to be known with certainty. Stochastic models on the other hand incorporate probability distributions for one or more variables, or uncertain error terms in relationships between variables. Thus a forestry bioeconomic model could take account of variable rainfall, crop hazards (e.g. pests, wildfires) and timber prices. With a stochastic model, the measure of performance could include both expected payoff and a measure of uncertainty. It is generally considered that a refinement in modelling introduces a need to incorporate uncertainty. That is, as the modeling is carried out in increasing detail, it becomes necessary to introduce of stochastic variables.

3. ELEMENTS OF MODELS

It is useful to introduce some further terminology to assist in the discussion of models. A convenient classification (drawing on Naylor *et al*. 1966) is to divide models into the following elements:

(i) *components* or building blocks. For example, in a forestry model these could include trees, soil and weather, or at a more aggregate level nursery, plantation stands, workforce, and machinery and equipment. By definition, management is outside but exerting planning and control functions over the system (and hence not a component of the model).

(ii) *variables*, the levels of which may be determined outside the system (exogenous) or within it (endogenous), or which may describe the state of the system (status variables). Exogenous variables may be under the control of management (*decision variables*) or not controllable (*environmental variables*). For a forestry operation, silvicultural treatments and associated labour and material inputs typically are controllable, while prices of labour and other inputs and of products generally are non-controllable. In simulation experiments, exogenous variables form the inputs which drive the model. Those variables under the control of management form the decision or policy instruments, the levels of which are adjusted in simulation experiments. Non-controllable exogenous variables may be modelled as either fixed values (or predetermined time series) or as probability distributions from which random values are generated. The desirability of particular management policies is assessed in terms of one or more model outputs or levels of endogenous *performance* variables. In a forestry model, these might include, mean annual increment (MAI), and NPV from a plantation. *Status* variables record the state of the system at each period in time. Examples for a plantation include mean tree height and girth at various ages, and also financial variables such as cash and debts. The growth rate of trees will depend on growth in previous time periods, sometimes referred to as a ‘feedback loop’.

(iii) *functional relationships*. These indicate how variables are interrelated, e.g. by linear or non-linear equations, and with or without time lags. Included are *identities*, which are true by definition (e.g. value = price x quantity sold) and *operating characteristics* which have to be estimated statistically or subjectively. The latter might include the relationship between tree height and age. Functional relationships give the system its unique behaviour, and needless to say the reliability of any systems model depends vitally on how accurately the relationships are identified and estimated.

(iv) *parameters*: these are the coefficients of the operating characteristics, values of which can only by estimated to within given confidence levels.

Any systems model may be summarized by the following symbolic relationship:

$$Z = f(X,Y,S,A)$$

where Z is a set of performance variables
X is a set of policy variables
Y is a set of environmental variables
S is a set of initial levels or status variables (including initial resource levels)
A is a set of parameter values, and
f signifies that a functional relationship exists between the variables in the various sets (i.e. f represents the model).

4. THE STEPS IN A SIMULATION STUDY

Carrying out a simulation study is like carrying out any other quantitative study (though sometimes a bit more complicated). The so-called *scientific method* is employed, which means a series of steps are performed in a logical fashion to achieve the overall task. The terminology for simulation steps varies between experts, but the following is a workable classification.

(i) identification of the problem
(ii) analysis of the system
(iii) synthesis of the model
(iv) programming the model to a computer
(v) testing the model
(vi) experimentation with the model
(vii) interpretation of results, and reporting to the relevant authority.

These steps are performed in the sequence listed, except that there is usually some cycling between them, e.g. testing the model may reveal the need to modify the structure then revise the computer program. Each of the individual steps will now be discussed briefly.

Identification of the problem

It is most important to identify clearly the study objectives in terms of the research or managerial problem which is to be examined. The nature of the model to be developed will depend on the problem to be analysed. Is the objective to understand the system or to prescribe management policies? If the latter, who is responsible for the system, and what are their goals, and what is wrong with present policies?

Analysis of the system

Once the problem is identified, the ‘systems analysis’ stage can be performed, in which the boundaries of the system, the relevant variables and their interrelationships are identified. This may involve drawing various charts or diagrams of the system.

Systems synthesis

The next step, of ‘systems synthesis’, consists of expressing the relationships between variables in symbolic form, and estimating the parameters of these relationships. Because of the high degree of flexibility possible, it is difficult to lay down rules for this major step, though some guidelines can be given. Where possible, statistical techniques should be used to estimate relationships between variables. Distributions or random variables can be obtained by testing the goodness-of-fit of alternative probability models using say the chi-squared test. Where historical data are scarce or are not considered relevant to future behaviour of the system, subjective estimation by people regarded as having *expert knowledge* about the system may be preferable. It is usually recommended to start with a relatively simple model, and gradually extend and refine it. To the extent that sub-systems are sufficiently independent, the model should be constructed in the form of a number of relatively self-contained *modules*, allowing these to be programmed and tested separately. Existing models of similar systems should be examined for relevance, since it may be possible to obtain ideas or even adapt modules from them.

Programming to a computer

Once a prototype version of the model is constructed, computer programming can commence, using a spreadsheet package or – if the model is to complex for this –

using a computer programming language such as Visual Basic, FORTRAN or C, or a specialist simulation modelling language such as Simile.

Testing the model

Having constructed a working model, it is necessary to test whether this model is an adequate representation of the real system. There is far from general agreement on how systems models should be tested, but a workable approach is to divide testing into *verification*, *validation* and *sensitivity analysis*. Verification is the process of testing whether the model takes its intended structure, i.e. whether the model is free of logical errors and whether the computer program performs as intended. Validation examines the broader question of whether the intended structure truely represents the real system. Validation efforts often lead to further refinement of the model.

Once a model has been validated as far as practicable, the effect of remaining errors on parameter estimates may be assessed through sensitivity analysis. If the purpose of the model is to identify optimal management policies, errors in estimates of performance due to inaccurate parameter values may not be of concern unless they lead to identification of inferior policies as optimal. That is, we are not concerned with the predictive ability of the model in absolute terms so much as the model's ability to correctly rank alternative management policies. For this reason, it is desirable to include a sensitivity analysis with respect to optimal values of decision variables. Sensitivity analysis usually involves adjusting parameter values by small amounts, and calculating various sensitivity criteria. Sensitivity may be expressed quantitatively in terms of 'elasticity' of performance with respect to parameter levels, or elasticity of optimal management policies with respect to parameter values. High sensitivities (elasticities) give cause for concern about the reliability of a model.

Performing experiments

Once sufficient confidence has been gained in a model, a variety of simulation experiments may be conducted. Various designs can be used for these experiments, as discussed later in the module. Many of the principles of experimental design as applied in the physical sciences are relevant to simulation experiments. Management policies may be specified in terms of a single policy variable or a number of variables. In practice, optimal (or at least desirable) levels of a number of policy variables often have to be determined simultaneously. For example, a forest manager may wish to know what fertilizer strategy, pruning and thinning regime and harvest schedule maximizes returns from a plantation. In the language of experimental design, the decision variables which are identified are known as *experimental factors*, and any combination of levels of these factors (i.e. any management policy) is known as a *treatment*. Measures of performance of the system are known as *response* variables. A computer run in which a number of treatments are evaluated is known as a simulation *experiment*. If random variability is built into the model (i.e. if the model is stochastic), then it is necessary to evaluate each treatment or policy under a number of different environments, i.e. to include *replication* treatments in the simulation experiment.

Experiments conducted on a computer also have important differences from real-world experiments. Three main sources of difference arise:

(i) compression of time. Because of the speed of computing and the low cost of computer time, it is usually possible to include a larger number of treatments and a greater degree of replication.

(ii) sequential processing. Traditionally, each of the treatments in a real-system experiment is evaluated at the same time. For example, in a plantation fertilizer experiment, the complete experimental design is decided, with all plots planted at the same time, and fertilizer applied to each on the same day, and girth and height meansurements made at the same times. On the other hand, because a computer is a sequential processor, treatments are evaluated sequentially in a computer simulation experiment. This means that the performance level for the first treatment is known before the second treatment is evaluated, and the performance for the first and second treatments are known before evaluating the third, and so on. Sequential processing opens the opportunity to use information gained from earlier treatments to direct factor levels in later treatments, within the same experiment. 'Optimum-seeking' experimental designs are discussed later in the module.

(iii) control over experimental variability. In a computer simulation experiment, the variability in the 'environment' is under the control of the researcher. A *random number generator* is used to produce numbers between zero and one, and these are transformed to random observations from the distributions specified for the random variables. If the random number generator is given the same *seed* for each treatment then the treatments are evaluated under the same sequences of random numbers, i.e. under the same environments. This reduces random variability in response levels between treatments relative to independent seeding (i.e. not re-seeding the random number generator). The result is greater power to detect differences between treatments for a given sample size (number of replicates).

Analysis and interpretation of computer output

Simulation experiments often generate reams of computer printout, and this output must be distilled and interpreted to a form useable by managers in a decision-support role. Since experiments conducted using stochastic simulation models do not provide exact results, estimated performance levels should be thought of in a confidence interval context.

5. EXAMPLE OF AN INVENTORY SIMULATION

Some of the concepts introduced above will now be demonstrated with reference to a simple inventory model.

Example 1

A seedling nursery faces uncertain demand for Mahogany seedlings. Recent experience suggests that demand can be approximated by a uniform distribution with a range of 5,000 to 10,000 seedlings in autumn, and 3,000 to 5,000 seedlings in each of winter, spring and summer. Seedling production cost is $300/1000 seedlings and the sale price is 60c/seedling. Seedlings are grown to be ready for planting in autumn, but may be retained until the following summer (after which they must be discarded), the holding cost being $120/1000 seedlings for each quarter.

Simulate quarterly marketing of Mahogany seedlings over four years, for a policy of growing 20,000 seedlings per year.

The spreadsheet model

The simulation model may be developed as an Excel spreadsheet, as in Table 1. The level of the decision variable (autumn inventory level) and parameter values (prices, costs, demand limits) are listed at the top of this sheet. For convenience, the year runs from autumn (the preferred planting time) through to summer. Each quarter is represented by a row in the simulation, and an estimate of net revenue is obtained for each quarter. Quarterly demand is obtained by the formula

lower demand level + random number x (demand range)

where the random number is from a uniform distribution in the range zero to one, obtained by the spreadsheet function RAND() (see Appendix A). Quarterly sales are obtained as

MIN(inventory level, demand level)

and any quantity left unsold is transferred to inventory in the following quarter, except that no inventory is carried forward from summer. Quarterly sales revenue is obtained as sales quantity multiplied by price net of production cost. Quarterly net revenue (or loss) is obtained as sales revenue less holding costs. Average annual revenue is obtained by summing quarterly net revenues over the 16 quarters and dividing by four.

In this example, the mean annual net revenue is about $3,980. Various inventory production policies (treatments) could be compared, by changing the 'Annual production' level, and re-running the simulation to determine the mean annual net revenue. In this way, the optimal production level could be determined. Because of the random element in demand, it would be preferable to increase the number of years over which sales are simulated, relative to the four years in this illustration.

While While separate random numbers would be generated for each inventory policy simulated, leading of confounding of effects of inventory policy and demand environment, this can be overcome by 'freezing' the sequence of random numbers.

Table 1. Spreadsheet model for inventory simulation example

Parameters of the simulation model

Production cost ($/1000)	300
Sale price ($/seedling)	0.6
Holding cost ($/1000 seedlings/quarter)	80
Autumn demand - lower limit (1000)	5
Autumn demand - upper limit (1000)	10
Off-season demand - lower limit (1000)	3
Off-season demand - upper limit (1000)	5

Simulation of quarterly sales and net revenue

Year	Quarter	Inventory (1000)	Random number	Demand (1000)	Sales (1000)	Holding cost ($)	Residue (1000)	Sales rev. ($)	Net rev ($)
1	Autumn	20.00	0.8501	9.25	9.25	0	10.75	2775	2775
1	Winter	10.75	0.4579	3.92	3.92	859.94	6.83	1175	315
1	Spring	6.83	0.6244	4.25	4.25	546.67	2.58	1275	728
1	Summer	2.58	0.3745	3.75	2.58	206.77	0.00	775	569
2	Autumn	20.00	0.3107	6.55	6.55	0	13.45	1966	1966
2	Winter	13.45	0.5514	4.10	4.10	1075.72	9.34	1231	155
2	Spring	9.34	0.5205	4.04	4.04	747.50	5.30	1212	465
2	Summer	5.30	0.4510	3.90	3.90	424.21	1.40	1171	746
3	Autumn	20.00	0.2979	6.49	6.49	0	13.51	1947	1947
3	Winter	13.51	0.6859	4.37	4.37	1080.85	9.14	1312	231
3	Spring	9.14	0.6435	4.29	4.29	731.10	4.85	1286	555
3	Summer	4.85	0.8500	4.70	4.70	388.14	0.15	1410	1022
4	Autumn	20.00	0.7927	8.96	8.96	0	11.04	2689	2689
4	Winter	11.04	0.8044	4.61	4.61	882.90	6.43	1383	500
4	Spring	6.43	0.7076	4.42	4.42	514.19	2.01	1325	810
4	Summer	2.01	0.3933	3.79	2.01	160.97	0.00	604	443
Mean annual net revenue									3979

6. VALIDATION OF SIMULATION MODELS

As noted above, validation is a critical task in development of a simulation model. While validation may be applied to the assumptions of the model, and to the various sub-models, in practice tests are usually concerned with the reasonableness of outputs or predictions of the overall model. In this regard, it is often recommended that statistical tests be used to compare output of the model with that of the real system, where both have been generated under the same management policies and environments. These tests examine hypotheses of the general form

H_0: outputs of the real system conform to those of the model
H_1: outputs of the real system differ from those of the model.

The output of complex (dynamic, stochastic) simulation models typically consist of time series for performance, which can include various statistical contributions (e.g. trend, seasonal effects). Since the two series of outputs may conform with respect to some properties and differ with respect to others, a comparison needs to be made of the various parameters (e.g. means, variances, autocorrelations and trends), and of overall distributions. On the face of it, these tests should allow a thorough check of all the statistical properties of the two output series. However, in practice a number of problems arise with statistical validation.

Often, little output is available from the real system with which to compare output from a model. All available data tend to be used in model construction, yet it is desirable for the data for testing purposes to be independent of that used when constructing the model.

Important differences arise between traditional applications of hypothesis tests and their role in validation of systems models. The motive behind traditional statistical testing is usually to demonstrate that a particular null hypothesis is false; the null hypothesis is simply set up as a 'straw man'. For example, when comparing two sample means the null hypothesis may state that the two underlying population means are equal, while the alternative hypothesis may state that they are unequal, i.e.

H_0: $\mu_1 = \mu_2$ (mean predicted output = mean real system output)

H_1: $\mu_1 \neq \mu_2$

and it may be possible to 'prove' that H_0 is false in the sense of demonstrating that the probability of μ_1 equalling μ_2 is negligible. When applying a validation test, the motive is usually to accredit the model, i.e. prove H_0 is true. Comparison of the t-statistic with critical values provides an indication of the probability that H_0 is false. If this probability is low, the model can be declared invalid. But if the probability is high (i.e. above the chosen significance level) then all that may be concluded is that the model has not been demonstrated to be invalid, beyond reasonable doubt. To state that the model must therefore be valid is at best a tenuous statistical inference.

Traditional hypothesis testing is designed to limit the frequency of type 1 errors. The cost of a type 1 error in terms of declaring a valid model as invalid may not be great; unnecessary marginal refinements may be made to the model. On the other hand, the cost of accepting as valid an invalid model may be substantial, both in terms of loss of credibility of the modeller if his or her creation is later found to be misleading, and in dollar terms for users from making incorrect management decisions based on information generated by the model. For a given sample size, the lower the significance level the greater the incidence of type 2 errors (accepting as valid models which are in fact invalid). In other words, to make the test procedure more stringent in terms of ability to reject invalid models, it is necessary to increase the significance level, say to 20%.

A further problem with statistical validation is that the assumptions underlying the tests may not be appropriate. For example, consider the t-test on mean model and real-system outputs. A paired-sample *t*-test is likely to be more appropriate than a test of independent samples, when the outputs have been generated under the same environments and managements. This paired-sample test assumes that successive differences are independent; in reality output series are often positively autocorrelated over time, with the result that the variance of differences between model and real-system outputs may be seriously underestimated. A modified form of *t*-test may be used which takes account of autocorrelations between paired differences for various time lags. Even this test requires the (relatively weak) assumption of second-order stationarity in differences, i.e. the correlation between differences depends only on the number of time periods between them, and is constant over time.

As illustrated by the above discussion, the application of traditional statistical tests to model validation is not a simple matter. It is no wonder that most model testing has been subjective in nature, such as graphical comparison of time series of model and real-system output, and assessment of the reasonableness of model output by people judged as experts with respect to the system.

Another reason for concern about statistical validation procedures is that they are being applied to a moving target. Simulation models typically evolve over time. A need for a model is identified, and a prototype model is developed and implemented to fulfil this need. Often, limitations of the model are recognized, and the model is further refined. As well as applying the model for its initial design purposes, other uses are often discovered, and extensions of adaptions of the model are undertaken to meet these other purposes. As a consequence, it is generally conceded that there is no clear finish to the testing of a systems model. Rather, confidence is gradually built up in a model over time as tests are performed, new information is obtained, and new versions of the model are produced.

7. THE DESIGN OF SIMULATION EXPERIMENTS

Suppose a simulation experiment involves two decision variables or experimental factors. Combining each level of one factor with each level of the other leads to *full factorial* design. The full factorial design is conceptually simple and well suited for simulation experiments in which there are only two or three factors, each taking only a few levels. In general, if there are *m* factors, each taking *n* levels, then there are *mn* distinct treatments. If *m* is more than two or three, the number of treatments can increase dramatically, and the experiment may become unmanageably large, even when the number of replicates is small. Alternatives to the full factorial design, such as the *fractional factorial* and *central composite* designs, allow for some reduction in treatment numbers for given m and n.

An alternative to designs which are specified in advance of the simulation experiment, is to include a set of rules in the computer program to select factor levels during the experiment on the basis of information gained from earlier treatments. The simplest of these designs is *steepest ascent* (or steepest descent). This is an iterative procedure in which the slope of the response surface is estimated with respect to each experimental factor, then all factor levels are adjusted so as to achieve the most rapid increase in performance. More sophisticated *hill climbing designs* may be adopted. These will involve programming the procedure for allocation of treatments on the basis of progressive performance during the experiment, or using sub-routines which have been developed by others for this purpose, which are available from a variety of sources.

Yet another alternative is *random search*, in which treatments are chosen simply by selecting the level of each factor at random. It is necessary to place upper and lower bounds on factor levels in order to confine the search area, based on prior knowledge of the system and exploratory computer runs. Factor levels are then sampled from uniform distributions over these ranges. If the number of treatments evaluated is sufficiently large, then there is a high probability that near-optimal policies will be identified. Often, computer output is obtained only for those treatments for which performance exceeds a specified threshold level. Random search is relatively easy to program on a computer. This experimental design procedure is particularly useful when the levels of policy variables are discrete, and cannot be approximated satisfactorily by continuous variables. This design is also used where activities form complementary or mutually exclusive sets. More sophisticated random search routines include provision for adjustment of probabilities or *heuristic learning* during the experiment; this may involve narrowing the search ranges or departing from uniform distributions.

8. SIMULATION IN PERSPECTIVE

Computer simulation relies on a overall systems philosophy, which asserts that components of a system should not be viewed in isolation and in a reductionist manner, but rather in terms of the complex interrelationships and interactions between variables. An attempt is made to analyse carefully all the components of the system and their interrelationships, and to represent these in a simplified and abstract model. This model is used to conduct experiments in which the behaviour of the system is simulated for various environments and management policies.

Simulation has proved to be a powerful and versatile approach to improving the understanding of systems and determining near-optimal management policies. A number of advantages are afforded over analytical techniques for problem solving, especially with regard to the handling of uncertainty, multiple goals, non-linear relationships and other real-world complexities. On the other hand, some difficulties arise in simulation studies. The successful application of this technique requires considerable experience with modelling, an understanding of statistical techniques and their limitations, and substantial human and computing resources. It is not usually possible to make use of a recognized model structure or existing computer program, and particular attention must be directed to validation of simulation models. A variety of designs may be exploited in simulation experiments, including optimum-seeking designs which take account of the special features of computer-based experimentation.

REFERENCES

Dayananda, D., Irons, R., Harrison, S., Herbohn, J. and Rowland, P. (2002), *Capital Budgeting: Financial Appraisal of Investment Projects*, Cambridge University Press, Cambridge.

Shannon, R.E. (1975), *Systems Simulation: the Art and the Science*, Prentice-Hall, Englewood Cliffs.

APPENDIX A: GENERATION OF RANDOM VARIATES

A critical step in any stochastic simulation model is to generate values of random (stochastic) variables or *variates*. A wide variety of forms of probability distributions has been recognized as reflecting the behaviour of particular random variables. The starting point for producing (or *generating*) values from any form of probability distribution is to use a random number from a known distribution. In practice, this is usually a number from a uniform distribution over the range zero to one, for example as provided in the Excel function '= RAND( )'.[4]

Generating uniform variates over a specified range

Perhaps the simplest form of *continuous* probability distribution to use in simulation models is the uniform distribution. Here *a* is the smallest value and *b* the largest value that the variable is expected to take. Since the area under any probability curve of function must be unity, and the range of values of the variable is *b-a*, the height of the curve must be *1/(b-a).*

A value y from a uniform distribution can be generated by taking a random number *r* then applying the expression

$$y = a + r(b-a),$$

where the random number is be obtained using the spreadsheet function '=RAND( )', i.e. the lower limit plus the random number times the distance between the upper and lower numbers. For the first demand estimate in Table 1, this becomes

$$y = 5 + 0.8501\ (10-5) = 5 + 4.2505 = 9.25$$

To obtain a series of values on the random variable, simply repeat this process a number of times (on a computer).

Generating values from a discrete distribution

In essence, this involves expressing the distribution in cumulative form, and then associating random numbers with cumulative probability ranges. For example, suppose timber price (in $\$/m^3$) can be

[4] The procedure used on the computer is typically to divide a very large number by another large divisor, and to take the remainder as a fraction of the divisor (hence yielding a number between zero and one). As well, the remainder is multiplied by a third large number, and the product divided by the original divisor. By repeating this process, a series of random numbers can be generated. The numbers are sometimes called 'pseudorandom' because given the same initial three large numbers (and same computer accuracy), the same series will always be produced.

represented by the discrete values 30, 40 and 55, with probabilities 0.3, 0.5 and 0.2. Cumulative probabilities are then 0.3, 0.8 and 1.0, and ranges of random numbers as assigned 0 to 0.3000, 0.3000 to 0.8000 and 0.8000 to 1.0000.

Now suppose the sequence of random numbers 0.2488, 0.8324, and so on is obtained. The first of these numbers falls in the first cumulative probability range hence a timber price of $\$30/m^3$ is generated. The second number falls in the range 0.8000 to 1.0000 hence a timber price of $\$55/m^3$ is generated. Proceeding in this way, a sequence of price 'observations' can be generated for which the relative frequencies approximate the discrete probabilities, providing the sample is sufficiently large.

Generating normal variates

The normal distribution is widely recognized in statistical methods as a commonly occurring or approximated distribution. There are several methods for generating random normal variates. The simplest is to take the sum of 12 random numbers from a uniform 0-1 distribution (e.g. a computer random number generator), subtract six, then multiply by the target standard deviation σ and add the target mean μ. The reason this works will not be explained here, but is associated with the Central Limit Theorem.

Generating values from a triangular distribution

A convenient distribution for fitting subjectively to random variables is the triangular distribution, defined in terms of the most pessimistic, most likely (modal) and most optimistic values. For example, for timber price these might be 30, 40 and 55 (in $\$/m^3$). If these points are called a, b and c, and a distance parameter is defined as

$$d = (b-a)/(c-a),$$

then using a random number r a value from this distribution y can be generated as:

if $r \leq d$ then $y = a + \sqrt{r\,(c-a)(b-a)}$

if $r > d$ then $y = c - \sqrt{(1-r)(c-a)(c-b)}.$

Procedures or 'recipes' can be developed along similar lines for sampling from a number of other continuous or discrete forms of probability distributions. Some computer packages in fact have these sampling procedures built in, such as the risk simulation package @RISK.